

AD-A251 861



2

UNDERSTANDING MECHANICAL SYSTEMS
THROUGH COMPUTER ANIMATION
AND KINEMATIC IMAGERY

FINAL REPORT

PATRICIA A. CARPENTER
MARCEL ADAM JUST
CARNEGIE MELLON UNIVERSITY

APRIL 1992

DTIC
ELECTE
JUN 09 1992
S A D

This report was prepared under the Navy Manpower, Personnel, and Training R&D Program of the Office of Naval Research under Contract No. N00014-89-J-1218. Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

This document has been approved
for public release and sale; its
distribution is unlimited.

92-14982



92 6 05 117

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; Distribution unlimited	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE				
4. PERFORMING ORGANIZATION REPORT NUMBER(S) ONR92-1			5. MONITORING ORGANIZATION REPORT NUMBER(S)	
6a. NAME OF PERFORMING ORGANIZATION Carnegie Mellon University		6b. OFFICE SYMBOL (if applicable)	7a. NAME OF MONITORING ORGANIZATION Cognitive Science Program Office of Naval Research (Code 1142CS)	
6c. ADDRESS (City, State, and ZIP Code) Department of Psychology Pittsburgh, PA 15213			7b. ADDRESS (City, State, and ZIP Code) 800 North Quincy Street Arlington, VA 22217-5000	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION		8b. OFFICE SYMBOL (if applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-89-J-1218	
8c. ADDRESS (City, State, and ZIP Code)			10. SOURCE OF FUNDING NUMBERS	
			PROGRAM ELEMENT NO. 0602233N	PROJECT NO. RM33M20
11. TITLE (Include Security Classification) Unclassified: Understanding Mechanical Systems Through Computer Animation and Kinematic Imagery.				
12. PERSONAL AUTHOR(S) Dr. Patricia A. Carpenter and Dr. Marcel Adam Just				
13a. TYPE OF REPORT FINAL		13b. TIME COVERED FROM 12-1-88 TO 1-30-91	14. DATE OF REPORT (Year, Month, Day) 92-04-30	15. PAGE COUNT 33
16. SUPPLEMENTARY NOTATION Supported by the Office of the Chief of Naval Research Manpower, Personnel, and Training R&D Program.				
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) Mechanical reasoning, Intelligence testing, Individual differences in problem solving.	
FIELD	GROUP	SUB-GROUP		
05	09			
19. ABSTRACT (Continue on reverse if necessary and identify by block number) One purpose of the research is to develop models of cognitive processes in understanding mechanical systems. A particular focus was on the processes in mentally animating the representation of a mechanical system, and the contribution of animation graphics in comprehension. Several studies, involving eye fixations, verbal protocols and process tracing, indicated that mental animation was difficult for individuals who were not knowledgeable about mechanics. Animation did help them determine the motion of individual components, but animation alone did not entirely compensate for the subject's difficulty in identifying relevant features and ignoring irrelevant features. A second goal of the research was to analyze the differences among individuals who are performing analytic reasoning tasks. The cognitive processes in a widely used, nonverbal test of analytic intelligence, the Raven Progressive Matrices Test were analyzed using experimental and modelling techniques. The processes that distinguish average and superior performance are the ability to induce abstract relations and the ability to dynamically manage a large set of problem solving goals in working memory.				
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Susan E. Chipman			22b. TELEPHONE (Include Area Code) (703) 696-4318	22c. OFFICE SYMBOL ONR 1142CS

INTRODUCTION

This is the final report for ONR contract NO0014-89-J-1218 between the Office of Naval Research (Manpower Committee) and Carnegie-Mellon University. Patricia A. Carpenter and Marcel Adam Just were the principle investigators. This contract covers work from December 1, 1988 through November 30, 1991.

The overall purpose of the research is to develop models of the cognitive thinking that constitutes understanding mechanical systems. The comprehension of mechanical devices, whether in preparation for operating, assembling, or repairing them, involves constructing a representation of the mechanical and physical properties of the device, including the motions and actions of the component parts and their dynamic interrelations. A particular focus of this research was on the processes that occur in mentally animating the representation of a mechanical system, and additionally, the processes in understanding animation graphics systems that display mechanical motion.

A second goal of this research is to analyze the differences among people who are good at various types of reasoning tasks and those who are not. Differences among individuals in their ability to reason is of obvious practical and scientific significance. An important facet of the completeness of a theory is to account not only for the effects of a task and situation, but also the systematic differential performance among individuals.

The potential applications for the Navy are most obvious in two areas. The first area is personnel selection, specifically, better interpretation of the processes that are assessed by existing achievement and skill tests as well as the potential for better design of future tests. The second area of potential application is in training. Animation graphics opens the possibility of new instructional techniques both in training and job situations. Research on the comprehension of such animation graphics has so far not kept pace with the rapid technological advances, so that relatively little is known about the cognitive processes that may make this technology more or less useful in training and learning situations.

The research approach is to develop fine-grained analyses of the reasoning and visual-perceptual processes in various types of problem solving tasks. The project utilizes data-intensive methodologies, such as eye fixations and verbal protocols, that allow us to monitor the cognitive processes on-line, as they occur, and to relate these measures to the eventual outcome, such as the correctness or type of response. Thus, these investigations seek to analyze the micro-structure of the problem solving processes, particularly in spatial and mechanical domains.

The following sections briefly summarize the research associated with the project and cites references to published descriptions of the work.



Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

1. Individual differences in working memory

The concept of analytic intelligence is a pervasive one in personnel selection, in psychometric theory and in testing more generally. In spite of the wide-spread use of "intelligence" tests, there is very little research on the actual processing that such tests evoke. In one line of research, we have pursued an analysis of the processes that occur during various cognitive tasks, such as spatial ability (Carpenter & Just, 1986; Just & Carpenter, 1985), verbal reasoning (Just & Carpenter, 1992; Carpenter & Just, 1989), mechanical problem solving (Hegarty, Just & Morrison, 1988; Just & Carpenter, 1987; Hegarty, Carpenter & Just, 1991), and complex reasoning (Carpenter, Just & Shell, 1990). The approach in all of these projects has been to use a variety of methods to analyze the ongoing thought processes of both more and less successful problem solvers, including eye fixations and "think aloud" protocols and other process-tracing methodologies (Just & Carpenter, 1976, 1988, 1987). These empirical studies are coordinated with the construction of detailed models of those processes, models that are often implemented as computer simulations. The scientific goal has been to combine a variety of techniques to specify the cognitive processes that underlie basic cognitive skills.

One series of studies has focused on characterizing reasoning, particularly focusing on the role of working memory. The initial research focused on a common psychometric test called the *Raven Progressive Matrices Test* (Raven, 1962). The Raven test, including the simpler Standard Progressive Matrices Test and the Coloured Progressive Matrices Test, is also widely used in both research and clinical settings. The test is used extensively by the military in several western countries (for example, see Belmont & Marolla, 1973). Also, because of its non-verbal format, it is a common research tool used with children, the elderly, and patient populations for whom it is desirable to minimize the processing of language. The wide usage means that there is a great deal of information about the performance profiles of various populations. But more importantly, it means that a cognitive analysis of the processes and structures that underlie performance has potential practical implications in the domains in which the test is used either for research or classification.

There are several reasons why the Raven test provides an appropriate test bed to analyze analytic intelligence. First, the size and stability of the individual differences that the test elicits, even among college students, suggest that the underlying differences in cognitive processes are susceptible to cognitive analysis. Second, the relatively large number of items on the test (36 problems) permits an adequate data base for the theoretical and experimental analyses of the problem-solving behavior. Third, the visual format of the problems makes it possible to exploit the fine-grained, process-tracing methodology afforded by eye fixation studies (Just & Carpenter, 1976). Finally, the correlation between Raven test scores and measures of intellectual achievement suggests that the underlying processes may be general, rather than specific to this one test (Court & Raven, 1982), although like most correlations, this one must be interpreted with caution.

Several different research approaches have converged on the conclusion that the Raven test measures processes that are central to analytic intelligence. Individual differences in the Raven correlate highly with those found in other complex, cognitive tests (see Jensen, 1987). The centrality of the Raven among psychometric tests is graphically illustrated in several nonmetric scaling studies that examined the interrelations among ability test scores obtained both from archival sources and more recently collected data (Snow, Kyllonen & Marshalek, 1984). The scaling solutions for the different data bases showed remarkably similar patterns. The Raven and other complex reasoning tests were at the

center of the solution. Simpler tests were located towards the periphery and they clustered according to their content, as shown in Figure 1. This particular scaling analysis is based on the results from a variety of cognitive tests given to 241 high school students (Marshalek, Lohman & Snow, 1983).

Insert Figure 1 - Marshalek et al. Results

Snow et al. also constructed an idealized space to summarize the results of their numerous scaling solutions, in which they placed the Raven test at the center, as shown in Figure 2. In this idealized solution, task complexity is maximal near the center and decreases outward toward the periphery. The tests in the annulus surrounding the Raven test involve abstract reasoning, induction of relations, and deduction. For tests of intermediate or low complexity only, there is a clustering as a function of the test content, with separate clusters for verbal, numerical and spatial tests. By contrast, the more complex tests of reasoning at the center of the space were highly intercorrelated in spite of differences in specific content.

Insert Figure 2 - Idealized Results

One of the sources of the Raven test's centrality, according to Marshalek, Lohman and Snow was that "... more complex tasks may require more involvement of executive assembly and control processes that structure and analyze the problem, assemble a strategy of attack on it, monitor the performance process, and adapt these strategies as performance proceeds..." (1983, p. 124). This theoretical interpretation is based on the outcome of the scaling studies. Our research also converges on the importance of executive processes, but the conclusions are derived from a process analysis of the Raven test.

A task analysis of the Raven Progressive Matrices Test suggests some of the cognitive processes that are likely to be implicated in solving the problems. The test consists of a set of visual analogy problems. Each problem consists of a 3 x 3 matrix, in which the bottom right entry is missing and must be selected from among eight response alternatives arranged below the matrix. Each entry typically contains one to five figural elements, such as geometric figures, lines, or background textures. The test instructions tell the test-taker to look across the rows and then look down the columns to determine the rules and then to use the rules to determine the missing entry. The problem in Figure 3 illustrates the format.

Insert Figure 3 - Sample Problem

The variation among the entries in a row and column of this problem can be described by three rules:

- *Rule A.* Each row contains three geometric figures (a diamond, a triangle and a square) distributed across its three entries.
- *Rule B.* Each row contains three textured lines (dark, striped and clear) distributed across its three entries.
- *Rule C.* The orientation of the lines is constant within a row, but varies between rows (vertical, horizontal, then oblique).

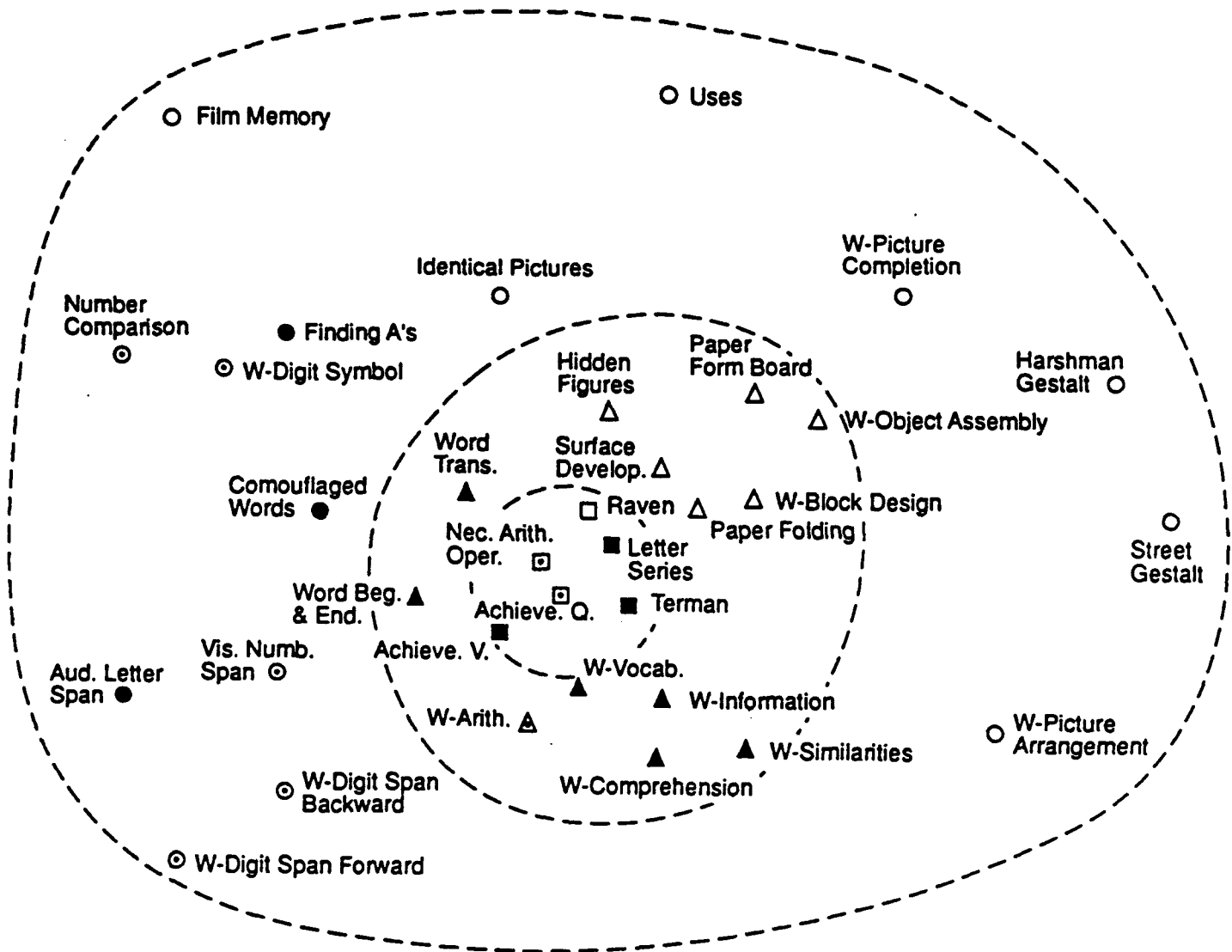


Figure 1

Figure 1. A nonmetric scaling of the intercorrelations among various ability tests, showing the centrality of the Raven (from Marshalek, Lohman & Snow, 1983, Figure 2, p. 122). The tests near the center of the space, such as the Raven and Letter Series Tests, are the most complex and share variance despite their differences in content (figural versus verbal). The outwardly radiating concentric circles indicate decreasing levels of test complexity. The shapes of the plotted points also denote test complexity: squares (most complex), triangles (intermediate complexity), and circles (least complex). The shading of the plotted points indicates the content of the test: white (figural), black (verbal) and dotted (numerical). (Reprinted by permission of authors.)

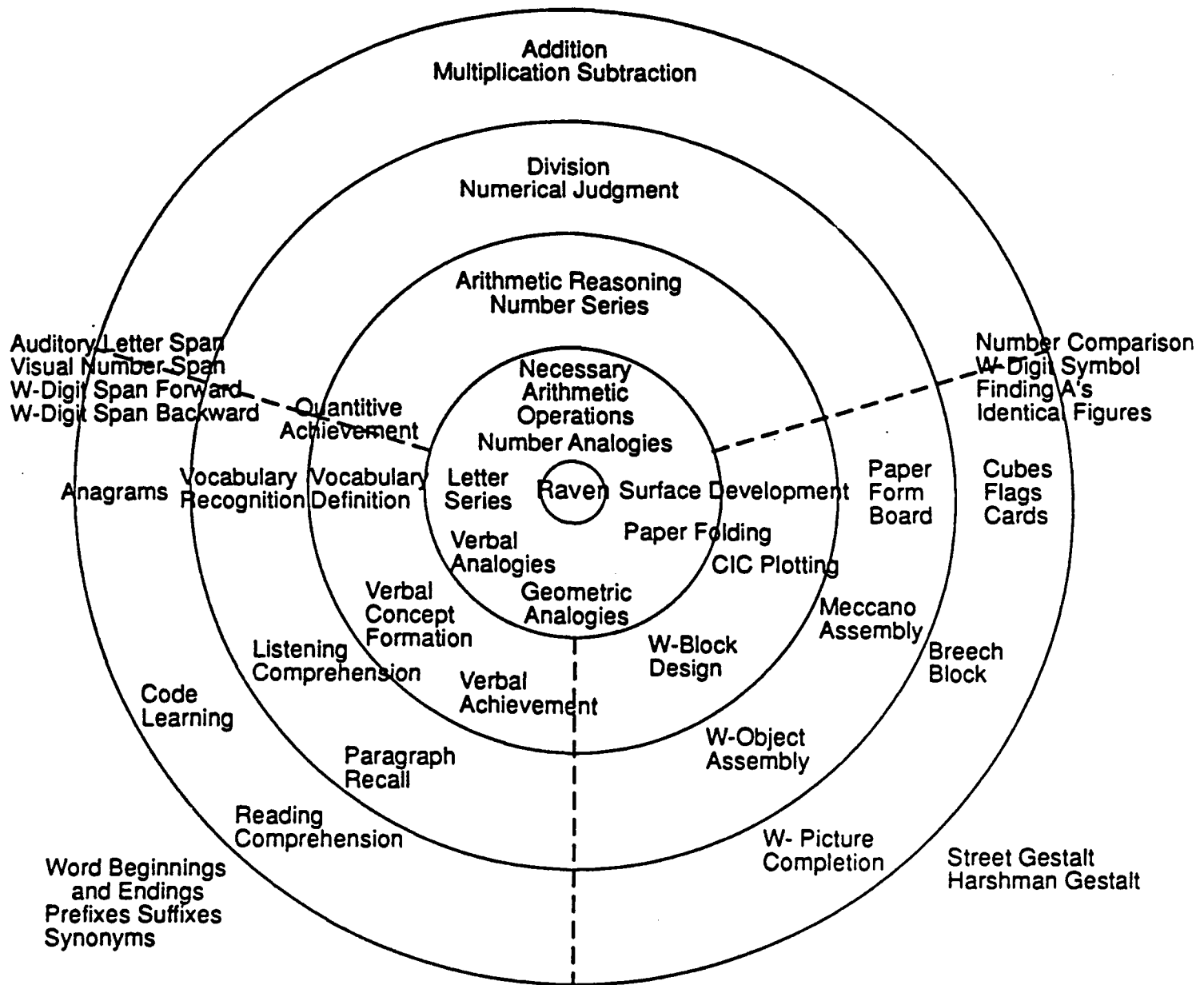


Figure 2

Figure 2. An idealized scaling solution that summarizes the relations among ability tests across several sets of data, illustrating the centrality of the Raven test (from Snow, Kyllonen & Marshalek, 1984; Figure 2.9, p. 92). The outwardly radiating concentric circles indicate decreasing levels of test complexity. Tests involving different content (figural, verbal, and numerical) are separated by dashed radial lines. (Reprinted by permission of authors and publisher).

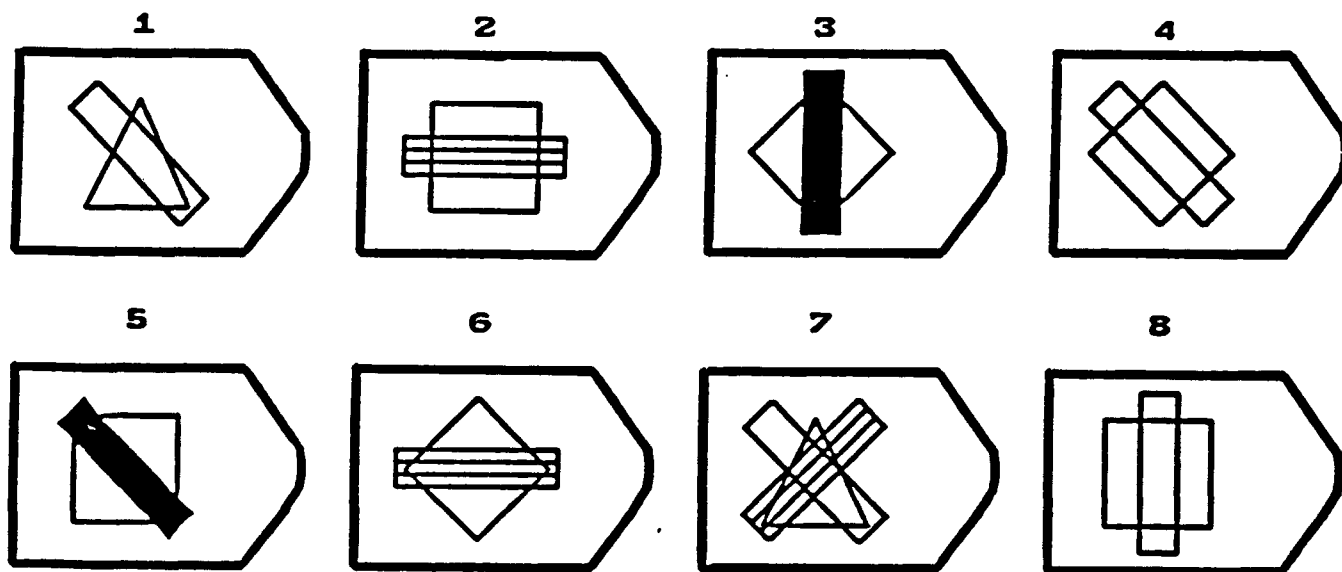
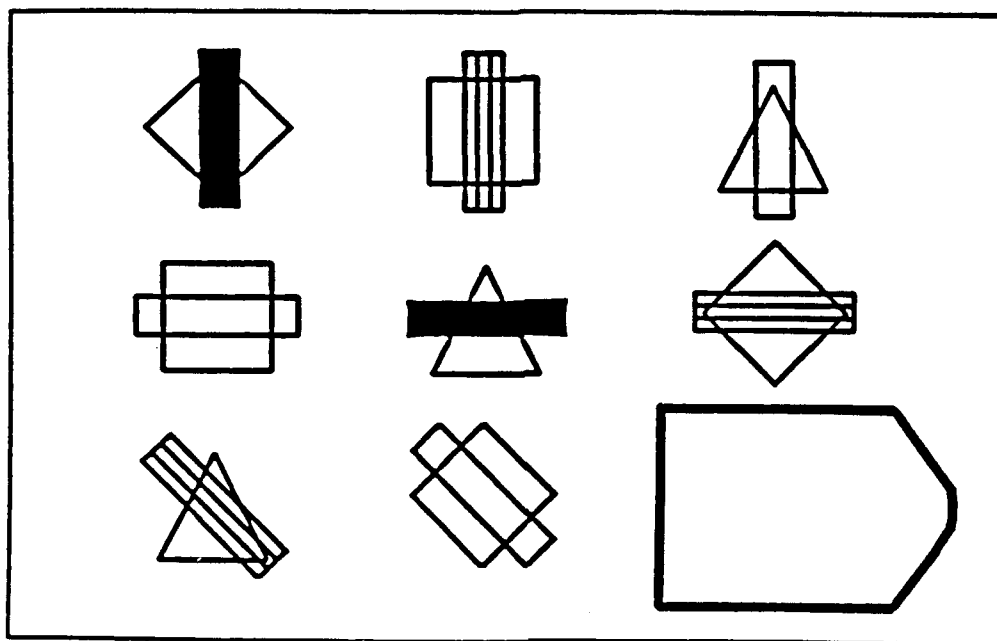


Figure 3

Figure 3. A problem to illustrate the format of the Raven items. The variation among the three geometric forms (diamond, square, triangle) and three textures of the line (dark, striped, clear) is each governed by a distribution-of-three-values rule. The orientation of the line is governed by a constant in a row rule. (The correct answer is #5).

The missing entry can be generated from these rules. Rule A specifies that the answer should contain a square (since the first two columns of the third row contain a triangle and diamond). Rule B specifies it should contain a dark line. Rule C specifies that the line orientation should be oblique, from upper left to lower right. These rules converge on the correct response alternative, #5. Some of the incorrect response alternatives are designed to satisfy an incomplete set of rules. For example, if a subject induced Rule A but not B or C he might choose alternative #2 or #8. Similarly, inducing Rule B but omitting A and C leads to alternative #3. This sample problem illustrates the general structure of the test problems, but corresponds to one of the easiest problems in the test. The more difficult problems entail more rules or more difficult rules, and more figural elements per entry.

The research is reported in Carpenter, Just & Shell (1990), which describes a theoretical model of the processes in solving the Raven test, contrasting the performance of college students who are less successful in solving the problems to those who are more successful. The model is based on multiple dependent measures, including verbal reports, eye fixations and patterns of errors on different types of problems. The experimental investigations led to the development of computer simulation models that test the sufficiency of our analysis. Two computer simulations, FAIRAVEN and BETTERAVEN, express the differences between good and extremely good performance on the test. FAIRAVEN performs like the median college student in our sample; BETTERAVEN performs like one of the very best. Figure 4 shows a flow-chart of the processes in BETTERAVEN.

The simulation had several modules (Figure 4) that encode the stimuli (symbolic descriptions of the figures), match the encoding to rules, generalize rules, and find the response. But the important part of the simulation that accounted for the difference between the median and best subjects was a goal manager. The goal manager kept track of multiple rules and allowed the system to backtrack in reformulating alternative rules. BETTERAVEN differs from FAIRAVEN in two major ways. BETTERAVEN has the ability to induce more abstract relations than FAIRAVEN. In addition, BETTERAVEN has the ability to manage a larger set of goals in working memory and hence can solve more complex problems. In a cognitive "lesioning" experiment, we changed the architecture of simulation to individual differences. We manipulated the capacity of the goal manager. This manipulation allowed the simulation to capture the differences between median and very best performing subjects.

Insert Figure 4 - BETTERAVEN

The contrast between the models specifies the nature of the analytic intelligence required to perform the test and the nature of individual differences in this type of intelligence. The processing characteristic that is common to all subjects is an incremental, re-iterative strategy for encoding and inducing the regularities in each problem. Thus, the paper argues that the processes that distinguish among individuals are primarily the ability to induce abstract relations and the ability to dynamically manage a large set of problem-solving goals in working memory.

Our current conception of working memory capacity is in terms of the amount of activation available for both maintaining and manipulating symbolic information in reasoning tasks. We have developed an interpreter for a production system architecture that can be set to have different amounts of activation (high amounts correspond to good ability). We can also use this simulation to investigate different strategies for what occurs to information

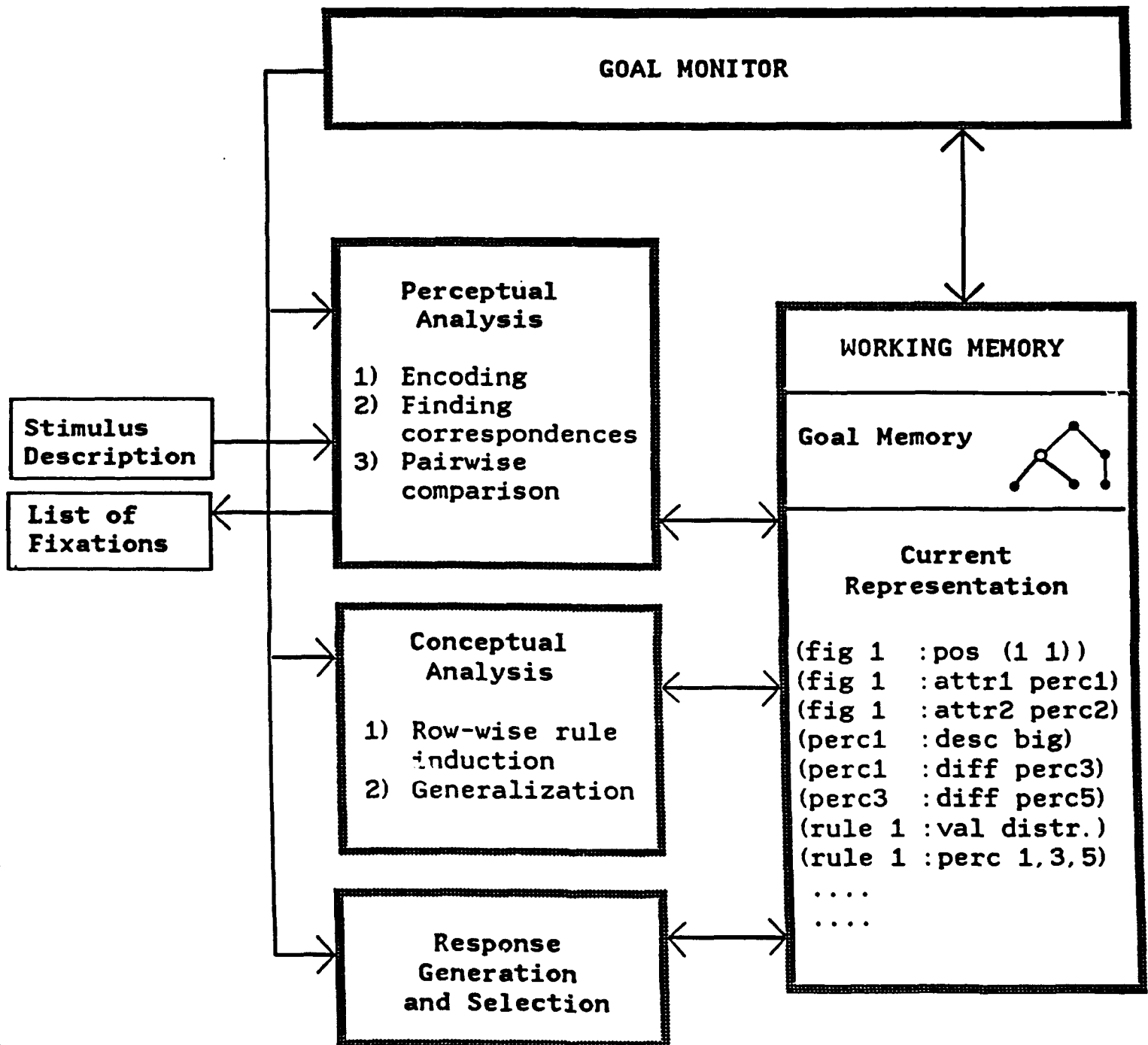


Figure 4

Figure 4. A block diagram of BETTERAVEN. The distinction from FAIRAVEN visible from the block diagram is the inclusion of a goal monitor that generates and keeps track of progress in a goal tree.

(forgetting plus slowing down of processing) when there is too little activation. This architecture has been applied to account for several types of systematic individual differences in language comprehension tasks; the architecture and the empirical results are described in detail in several recent publications (Carpenter & Just, 1989; Just & Carpenter, 1992; MacDonald, Just & Carpenter, 1992)

One conclusion of the research on individual differences in reasoning tasks is that a key determinant of performance in complex reasoning tasks is the availability of adequate working memory resources both for computing and storing intermediate goals and products during problem solving. In particular, the executive processes that enabled problem solvers to generate subgoals in working memory, to record the attainment of subgoals, and to set new subgoals as others were attained were critical to problem solving success and a source of individual differences. The executive processes were examined in studies of both cognitive processes and individual differences as determined by the Raven Progressive Matrices test; the latter is a measure of fluid reasoning ability and it typically correlates highly with complex visual problem solving.

Summary. This research suggests a very clear hypothesis about the nature of individual differences and task variation, more generally, in analytic problem solving. Ongoing research seeks to re-examine conceptions of spatial problem solving skill in light of this theoretical model of the constraints on analytic problem solving.

II. Mental animation and computer animation

As background, it is useful to remember that Navy training and maintenance manuals include diagrams with accompanying texts that are very complex for individuals who are less mechanically knowledgeable. The complexity of such material is illustrated in a typical excerpt taken from the Navy's book "Basic Machines and How They Work."

Insert Figure 5 - Navy Manual Excerpt

Our research has examined the processes used in interpreting such diagrams (and texts) and ways to use computer technology to impact on the comprehension of such materials.

Individual differences in these tasks were assessed by a common test of mechanical knowledge called the Bennett Mechanical Comprehension Test (1969), which has some items that are similar to those in the ASVAP. Typically, the item shows a mechanical situation and asks about some physical property (such as mechanical advantage) that does not require complex calculation. This isomorph of an actual item asks about the relative mechanical advantage of two systems. What is important is that it implicitly pits a relevant feature (the weights of the two objects) against an irrelevant feature (their distances from the source of the force -- the man). Less mechanically-experienced subjects and those who haven't had formal physics instructions are more likely to be misled by the distance factor. Their implicit model of the problem is that force flows from the source (the man) to the goal and so the first weight (answer B) will be lifted first. By contrast, the correct analysis is that the tension is equal throughout the rope and so the lighter weight (answer A) will be lifted before the heavier weight. Hence, the correct answer is A.

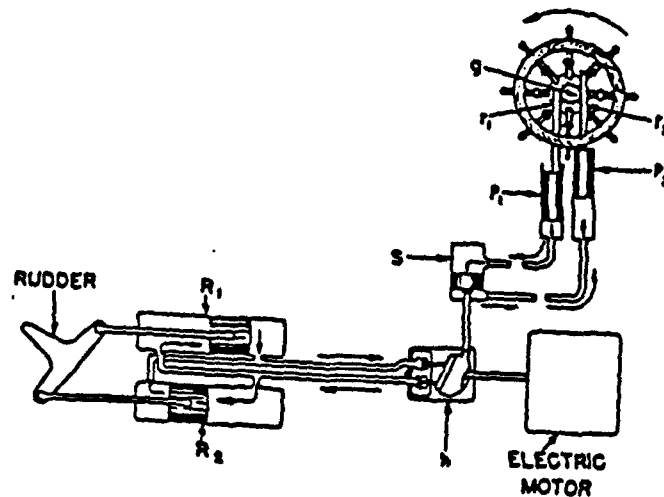
Sample description from Basic machines and how they work

Figure 10-11.—Steering is easy with this machine.

Hydraulics Aid the Helmsman

You've probably seen the helmsman swing a ship weighing thousands of tons about as easily as you turn your car. No, he's not a superman. He does it with machines.

Many of these machines are hydraulic. There are several types of hydraulic and electro-hydraulic steering mechanisms, but the simplified diagram in figure 10-11 will help you to understand the general principles of their operation. As the hand steering wheel is turned in a counterclockwise direction, its motion turns the pinion gear *g*. This causes the left-hand rack *r1* to move downward, and the right-hand rack *r2* to move upward. Notice that each rack is attached to a piston *p1* or *p2*. The downward motion of rack *r1* moves piston *p1* downward in its cylinder and pushes the oil out of that cylinder through the line. At the same time, piston *p2* moves upward and pulls oil from the right-hand line into the right-hand cylinder.

If you follow these two lines, you see that they enter a hydraulic cylinder *S*—one line entering above and one below the single piston in that cylinder. In the direction of the oil flow in the diagram, this piston and the attached plunger are pushed down toward the hydraulic pump *h*. So far, in this operation, you have used hand power to develop enough oil pressure to move the control plunger attached to the hydraulic pump. At this point an electric motor takes over and drives the pump *h*.

Oil is pumped under pressure to the two big steering rams *R1* and *R2*. You can see that the pistons in these rams are connected directly to the rudder crosshead which controls the position of the rudder. With the pump operating in the direction shown, the ship's rudder is thrown to the left, and the bow will swing to port. This operation demonstrates how a small force applied on the steering wheel sets in motion a series of operations which result in a force of thousands of pounds.

Figure 5

Figure 5. An example of text and a mechanical diagram from the Navy manual entitled Basic Machines and How They Work.

Insert Figure 6 - Mechanical Knowledge Test Item

It is reasonable to claim that people who understand mechanical systems can infer the principles of operation of an unfamiliar device from their knowledge of the device's components and their mechanical interactions. Individuals vary considerably in their ability to make this type of inference. A research project, reported in Hegarty, Just & Morrison, (1988), describes studies of performance of college students in psychometric tests of mechanical ability. Based on subjects' retrospective protocols and response patterns, it was possible to identify rules of mechanical reasoning that accounted for the performance of subjects of different levels of mechanical ability. The rules are explicitly stated in a simulation model which demonstrates the sufficiency of the rules by producing the kinds of responses observed in the subjects. Three abilities are proposed as the sources of individual differences in performance:

(1) ability to correctly identify which attributes of a system are relevant to its mechanical function,

(2) knowledge of a general functional relation between the attribute and the outcome (in this case, mechanical advantage) and the ability to use rules or relation consistently,

and (3) ability to combine information about two or more relevant attributes, initially qualitatively and then, quantitatively.

A series of protocol studies using carefully constructed items revealed that mechanical knowledge contributes to problem solving in the domain of mechanics in two ways: by increasing the likelihood of identifying the relevant attributes of a system, and by providing qualitative and quantitative rules that related these attributes to mechanical advantage. Without the relevant mechanical knowledge, such devices were internally represented in a fragmentary and non-functional way.

Mechanical reasoning by students and professional mechanics. In the next section, we describe several studies of mechanical reasoning in students and professional mechanics. This research was actually preliminary to the simulation and experimental studies reported above. Their importance here is to support the claim that the reasoning processes reported above are fairly general, both across different populations and different types of reasoning tasks.

Both book learning and hands-on experience under the car hood may improve mechanical reasoning. The studies, because they are all correlational, are only suggestive; nevertheless, we examined the impact of either practical mechanical experience and formal training in physics principles (operationalized as 1 year or more of college physics) on performance in the Bennett. "Mechanical experience" was operationalized as the person's report of some specific categories of practical mechanical experience, such as fixing small appliances, such as a toaster or a lamp; assembling a mechanical object, such as a bicycle or wheel barrow; or participating in activities, such as car repair. Either no mechanical experience or very sporadic and superficial experience was considered as "No Reported Experience." The same classification was used in a second study in which subject were asked to "talk aloud" while solving the problems to allow us to analyze their processes. The test scores were similar for the two tasks, suggesting that talking aloud did not impact on the overall problem solving success. In addition, to examine the contribution of spatial training to mechanical problem solving, we recruited 14 architecture majors; these students

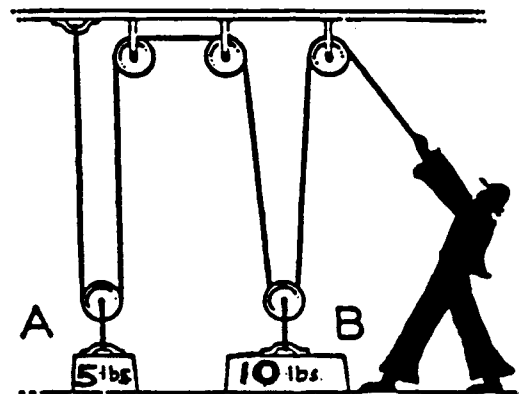


Figure 6

Figure 6. Example of an item that is similar to those in the Bennett Mechanical Comprehension Test. An irrelevant dimension (distance from the person) is pitted against a relevant dimension (weight of the items to be lifted). The correct answer is "A".

were essentially without formal physics training. The data indicate contributions of both mechanical experience and school courses for each group. The performance of the architecture students suggests that in spite of their lack of formal training relevant experience led to similar levels of performance.

Score (on 68 item Bennett Test)			
	Mechanical Experience		No Mechanical Experience
Unselected College Students -- Standard Test			
1 year physics	[n=11,13] 51.3 (s.d.=11)		42.8 (s.d.=12)
No physics	[n=4,21] 46.7 (s.d.=12)		33.0 (s.d.=8)
Unselected College Students -- "Thinking Aloud" Test			
1 year physics	[n=7,7] 53.5 (s.d.=5)		40.1 (s.d.=16)
No physics	[n=6,9] 37.5 (s.d.=15)		30.3 (s.d.=12)
Architecture Students			
No physics	[n=7,7] 51.1 (s.d.=8)		44.7 (s.d.=8)

This correlational analysis must be interpreted cautiously because of the obvious lack of control over the characteristics of who might end up in these various self-reported cells. Nevertheless, our data suggests that both mechanical experience and formal training are associated with higher scores. An additional point is that little of the formal instruction in college physics directly addresses the mechanical, electrical, and kinematic situations that are probed in the more practically-oriented items in the Bennett Test (1969). Consequently, the transfer that occurs from the course work may be at a more abstract level, such as learning the general principles. In addition, there is the fact that some difference in the scores may reflect more general subject selection characteristics of who takes college physics and who tends to have mechanical experience.

An additional point, which is relevant to the generality of our subsequent studies, is that the performance of the college students in most conditions is comparable to that cited in the Bennett Manual (1969) from a study of 315 applicants for "technical defense courses."

Mechanical Experience	No Reported Experience
[n=220,95] 41.7 (s.d.=8.6)	39.7 (s.d.=8.9)

These means from the manual are similar to those obtained in a much larger study reported of applicants for positions as firemen or policemen in New York City; the scores for the 879 high school graduates (removing data from those who had attended college) was

36.7 (s.d. = 9.7). Thus, the mechanical problem solving skills of most of the unselected subjects in our studies (with the exception of those who take college physics and report considerable mechanical experience) is roughly similar to that found in less selective populations.

It is also the case, however, that general problem solving skills confer some advantage in mechanical problem solving. The Bennett manual reports correlations between the Bennett and (an unspecified) intelligence tests of .40-.60. Consistent with the general result, in Experiment 2 (involving verbal protocols), the correlation for 29 subjects between Bennett and reported verbal SAT was .40. From this positive correlation, one might expect that more selective populations, selected by measures related to intelligence test scores, will tend to have higher scores on mechanical problem solving tests.

The important point here is that the processes in solving mechanical problems revealed in these students may generalize to other populations.

The naturally curious reader might wonder about the levels of performance by the professional mechanic -- the person to whom one entrusts one's Ford on the bad day that it stalls on Main Street. Are professional mechanics immune to the errors that plague mere mortals? In fact, some window on the extremes of experience was provided by a group of professional mechanics who solved the Bennett while talking aloud about their hypotheses and ideas. These were 13 adults who made their living as mechanics, including 3 airplane mechanics, 1 auto mechanic and 9 professional bicycle mechanics (two of whom had been mechanics in the Armed Forces). Their professional experience ranged from a minimum of 1 year to, at the other extreme, 28 and 41 years of experience (for two of the airplane mechanics). But "older" did not prove to necessarily be wiser; for these subjects, the correlation between years of professional mechanical experience and Bennett score was $r = .08$. Anecdotally, the actual mechanical experience differed among these individuals in spite of the shared job title. For example, the one auto mechanic said that most of his job was simply replacing parts that he was told to replace; he said that he seldom mechanically repaired broken parts. Perhaps not surprisingly, some mechanical jobs may not yield nuts-and-bolts experience that the layman naively associates with the position.

The overall scores of the group was 52, with an average of 15.9 errors out of the 68 problems. Interestingly, these professionals tended to make errors on the same problems that caused difficulty for the amateurs; the correlation over the 68 problems between the error rates for the two groups was .80. The reasons that mechanics gave for their answers were generally similar to those given by the other high scoring group -- college students who had at least 1 year of a college physics course. One major difference between the groups, summarized below, was that professional mechanics were more likely to not give a reason or simply restate the problem. It may be that professionals had implicit rules, but were less likely to have learned the explicit rule that college students could state in giving their rationale for an answer.

	Professional Mechanics	Experienced 1-yr Physics Students
No reason beyond problem statement	18%	5%
Mentions correct dimension and functional relation	43%	70%
Attend to an irrelevant dimension	17%	16%
Gives an incorrect rule or ignores the relevant dimension	23%	9%

In sum, successfully solving these mechanics problems involves learning the relevant dimensions, not being attracted by fortuitous variation in an irrelevant dimension; it also involves learning the general type of functional relation that links the relevant dimension to the issue (such as mechanical advantage). With formal training, students also learn precise quantitative rules and they may be more likely to learn the terminology to describe the relevant principles, even though quantitative rules are not required to solve the qualitative Bennett-type problems.

Comprehension of mechanical diagrams. The processes in successfully understanding a novel device or situation may seem complex, as witnessed by the difficulty that otherwise reasonable adults experience when confronted with the task of assembling a child's bike. Or, in the context of the Navy, consider the difficulty of understanding the explanation (given earlier) that we excerpted from the Navy's manual on how "hydraulics aids the helmsman". The nature and complexity of the processing in comprehending mechanical systems were apparent in a series of studies on how people reason about novel mechanical devices. One purpose of these studies was to understand the reasoning processes and sources of error; a related goal was to understand the role of mental animation and the depiction of animation in a graphics display. The question was whether a good graphics display could circumvent some of the difficulties that viewers have in understanding how mechanical things work.

In a typical experiment, the subject was shown a diagram and brief text that described a simple, novel device. The device was simple in the sense that it was created from a small number of common mechanical components, such as levers, gears, and ratchets. Although the device was similar, the task was not; many subjects had great difficulty figuring out what the device was doing. The difficulties experienced by these college students may be reasonably representative of the difficulties experienced by other, less selective adult populations.

The task that the subject faced can be understood by considering a typical device, called the ratchet device, shown in Figure 7. The task is to determine the motion of the wheel when the handle is pumped. [The answer is that the gear turns clockwise.] Figure 8 shows another example, called the pencil device. One can read the text, look at the diagram, and try to solve the problem given to the subject: The reader's task is to determine how the pencil moves when the drive gear moves clockwise? [Alternatively, the less mentally energetic might simply accept the answer that the pencil will trace a figure-8 that is oriented sideways.]

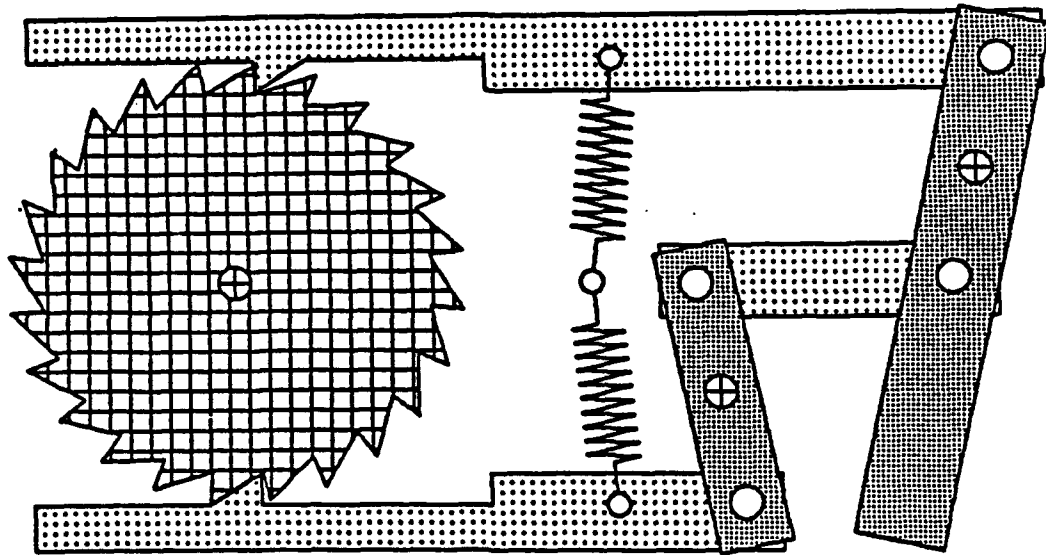
Insert Figure 7 and 8 - Mechanical Devices

To understand the intellectual sleuthing we undertook to pinpoint the difficulties in understanding how these things work, it is useful to describe the scene of the crime, so to speak; in this case, the scene was a study in which we recorded the eye fixations of subjects reading the text and inspecting the diagram. A task analysis of how subjects understand such a device suggested that comprehension involves determining how connected components interact and that this inference is done by "mentally animating" various joints along a line of action connecting the input force to the output. If so, then "mental animation" may be an important aspect of comprehension; more importantly, relieving the burden of mental animation by providing an animated display might improve the comprehensibility of such devices. Therefore, the research contrasts condition in which the display was static (as a diagram in a book) with one in which either the entire device or some component could be animated (usually at the viewer's discretion). The following sections describe three studies, one involving eye fixations, another with verbal protocols, and a third using a technology in which subjects explored the text and diagram by using a mouse to determine what components or sentences were visible. Throughout these studies, we found that for these devices, subjects who had more mechanical knowledge were not typically helped by the animation. It is as though they had sufficient schemas to infer the motions of the components and interactions for these devices. More surprisingly, the lower knowledge individuals were not helped very much either. The ability to animate the display decreased some of the their mistakes in mentally animating a joint; on the other hand, the difficulty of combining successive animations to determine interrelations among non-adjacent components appeared to be still problematic. "Seeing" the animated device is not a transparent perceptual process, but rather a complex cognitive perceptual process.

Experiment 1: Eye Fixations. In the first project, we analyzed how subjects inspected the diagram by recording their eye fixations. Forty undergraduates studied the ratchet device (after some preliminary familiarization with the procedure, display, and equipment). They were given as much time as they required. Then they were given 2-alternative and 4-alternative multiple choice questions about the functioning of the system, such as (1) What statement best describes the motion of the gear as the handle is pumped; (2) What happens to the small vertical connecting lever when the handle is pulled?; (3) What happens to the upper bar when the handle is pulled? Finally, they were asked to draw a picture of the device.

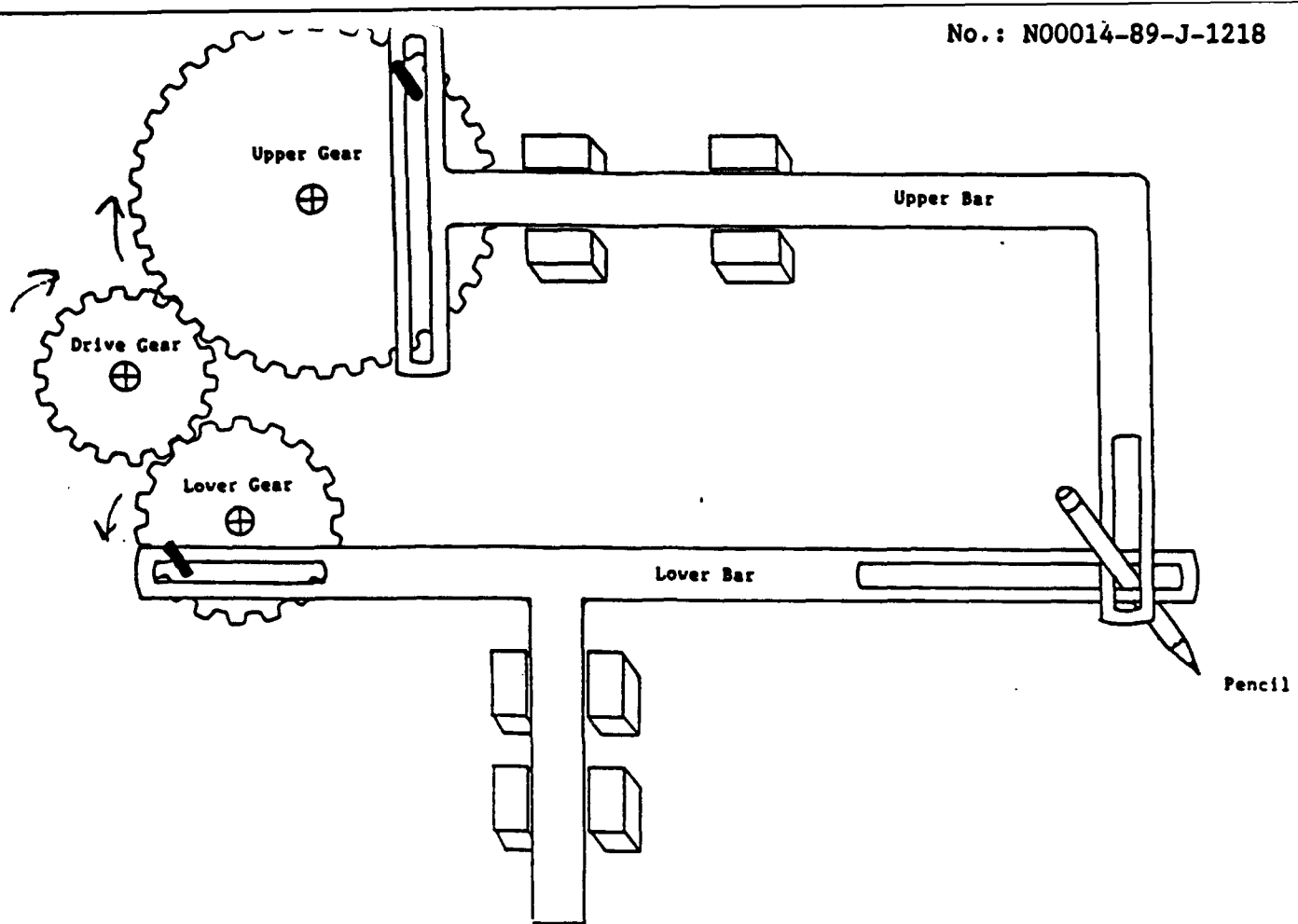
The subjects' mechanical knowledge was assessed by using a modified version of the Bennett in which we eliminated 20 questions that were least informative. The remaining 48 questions were those that had the best item-response characteristics (namely, an ogive function when the proportion correct for that item is plotted as a function of total score on the test) using data from the earlier studies of the Bennett Test. The subjects were divided into higher and lower scoring groups, with an average of 82% correct for the higher scoring groups and 61% for the lower scoring on the shortened Bennett.

The results. In answer to one of the major questions that motivated this study, animation had, at best, small and localized effects on subjects understanding of how the device worked. On the 9-question test asking about the device and its components, lower-knowledge subjects answered 2.7 and 3.8 questions correctly, and higher knowledge subjects answered 5.3 and 4.7 for the static and animated conditions, respectively; so that only knowledge and not animation had significant effects, $F(1,36) = 13.46$, $p < .01$.



This machine makes a gear wheel turn when the handle is pumped. The machine consists of a handle linked by a system of levers and bars to the gear wheel. When the handle is pulled, the upper bar turns the gear while the tooth in the lower bar slides over the gear teeth. When the handle is pushed, the lower bar turns the gear while the tooth on the upper bar slides over the gear teeth.

Figure 7



This machine moves a pencil when the leftmost gear, called the drive gear, is turned. The machine consists of an upper bar, a lower bar, a large upper gear, a smaller lower gear, and the drive gear, as labelled in the diagram. The upper and lower gears have pins mounted perpendicular to their surfaces and near their edges through which the gears interact with the bars. The pencil is perpendicular to the paper and mounted through both bars.

Your task is to figure out the shape of the line that would be drawn by the pencil when the drive gear is turned clockwise.

Figure 8

The most striking evidence of the fragmentary representation of lower-knowledge subjects was their subsequent drawings. We analyzed the drawings using a condition-blind scoring of the presence/absence of the major functional components; 70% of the lower-knowledge subjects' drawings had major errors, compared to only 45% for the higher-knowledge subjects. Moreover, the animated display did not ameliorate this difficulty; the likelihood of major structural errors was almost identical for the static and dynamic displays. Examples of the drawings (in Figure 9) most graphically convey their confusions and mistakes. As these samples indicate, many subjects, particularly low knowledge subjects, had fundamental misconceptions about the major functional components and their interrelation.

Insert Figure 9 - Drawings of Ratchet Device

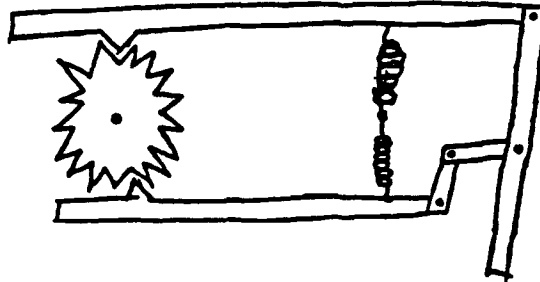
Lower knowledge subjects are more driven by the text in learning about the device, as indicated by relatively longer time (44 sec) they spent reading the text and smaller time (35 sec) inspecting the diagram than the higher knowledge subjects (34 sec and 40 sec, respectively), $F(1,36) = 8.53$, $p < .01$. Six seconds, on average, was spent in actually animating the display; this was additional time on the diagram, there was no influence of animation on the time spent reading the text. In spite of the reading and detailed inspection of the diagram lower-knowledge subjects had only fragmentary knowledge about the device.

Experiment 2: Verbal Protocols and Supplemented Descriptions. If lower knowledge subjects are so dependent on the text for guidance, perhaps a text that provided a great deal of guidance could break the bottleneck to improve their understanding. To test this hypothesis, we compared the standard description to a another version that was supplemented by instructions to imagine the motion of components in a sequence that corresponded with the line of action from input to output. In addition, we asked subjects to "think aloud" while they read the description and inspected the diagram. Forty students participated in the study, half of whom were given the supplemented description.

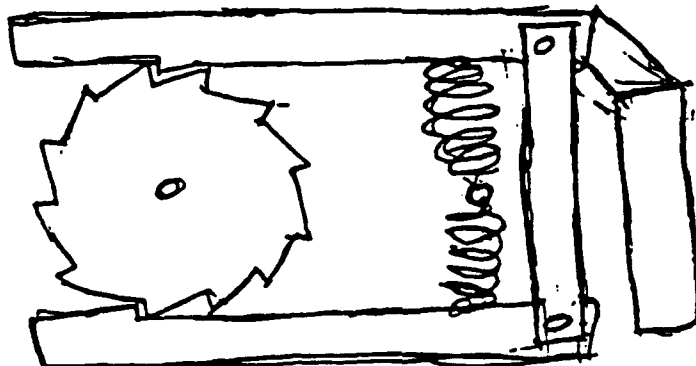
Disappointingly, the supplemented description was unable to break the bottleneck in comprehension. Few subjects (8 in the regular description condition and only 5 in the supplemented description condition) accurately described the motion of the gear wheel for the ratchet device. And overall, their question-answering skill was at a level similar to that in the eye fixation study. Some suggestion of the source of the difficulty came from the verbal protocols of subjects who failed to determine the motion of the gear. They were less likely to follow a lines of action; in addition, they were more likely to make an error in their inference about the direction of motion of a component. In sum, supplementary text did not improve comprehension, but the protocols strongly supported our task analysis that comprehension involved mentally animating the interacting components along a line of action. An inability to do such animation or follow a line of action was correlated with mistakes in understanding the device.

Experiment 3: Moving with a Mouse. The next hypothesis to be evaluated followed from the observation that better subjects mentally animate each joint as they follow a line of action; therefore, perhaps comprehension would improve if viewers were guided along lines of action and also were able to animate the display of a joint. Before describing the interesting technology that let us do this, it is useful to give the bottom line: Even this combination of animation and guidance did not dramatically improve the understanding of the lower-knowledge subjects. Subjects made fewer errors on the motions of individual

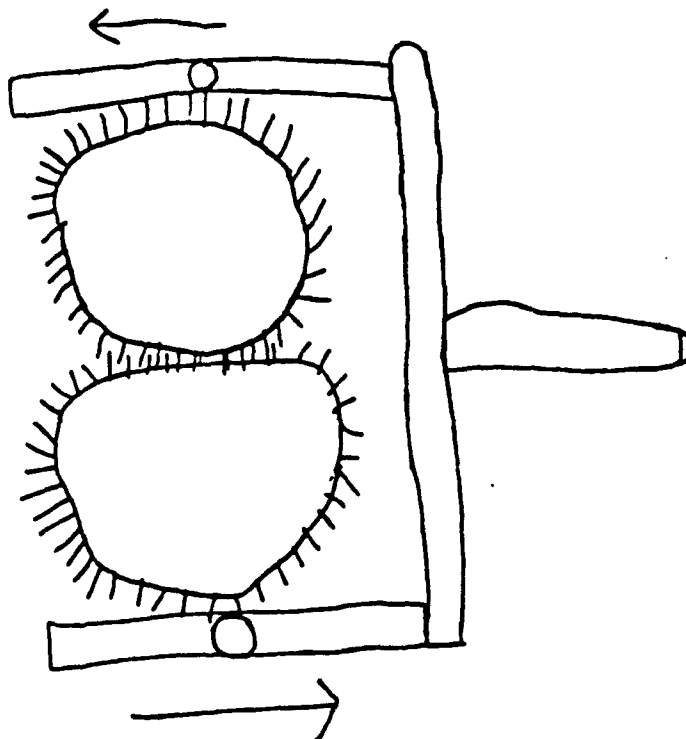
Subjects' Drawings.



Typical drawing by
Higher Knowledge
Subject



Typical drawing by
Lower Knowledge
Subject



Typical drawing by
Lower Knowledge
Subject

Figure 9

components, but they weren't helped on inferring interrelations or putting together successive components. To further jump ahead, it is possible to even speculate on the reasons why this intervention was unsuccessful. At this point, a plausible analogy might be made to educational research that attempted to develop reading skill in poorer skilled children (and adults) by trying to make them fixate at the same rate or in a similar pattern to good readers. The problem with this reading intervention is that it was aimed at an effect, not a cause. Good readers were faster as a result of better lexical, syntactic and semantic representations and processing, as well as more capacity to retain the intermediate and final products of their comprehension. The suggested analogy is that higher-knowledge individuals show the consistencies in tracing lines of action because they are more adept at accessing and assembling from their knowledge base appropriate representations that guide the encoding of relevant components, as well as their inferences about action.

The technology. The software was developed to be analogous to the "Moving window" technology used in reading. The idea is to limit what parts of the display are visually available and allow the subject to determine when and where to move to the next part. Thus, the experimenter can measure the sequence and duration for each portion of the text and diagram as they are viewed. Subjects selected which portion they saw by moving a mouse pointer into the region of the display screen associated with the portion. The amount of text visible in one portion was one paragraph. Hence, if a subject moved the mouse pointer onto some obscured text, all the words in that paragraph would become visible. Text was obscured by replacing every letter with an "x". For the diagram, either two or three continuous components were visible in a portion. (In the ratchet diagram, in addition to the two contiguous components that were displayed, the handle was also always visible to indicate whether it was in the push or pull phase of the cycle.) Device components were obscured by removing all detail, such as gear teeth, pivots and linkages, and replacing them with dimly illuminated blocks of grey. Consequently, the viewer always had some visual display of a component in their periphery, but no detailed information.

To ensure that subjects looked at components in the order specified in these texts, the control program would only permit subjects to select views in the same order as specified in the text. This program permitted subjects to select as many or as few components in a line of action as they chose, but the first component selected had to be the handle, and successive components had to follow the line of action.

To determine the effect of providing subjects with multiple views of the diagrams, an additional animation condition was run in which the entire display was visible. When the display was animated, all of the components of the device moved and were visible to the subject to inspect freely (as in the animation condition of the eye fixation study).

Subjects were also familiarized with real physical models before the experiment in order to ensure that difficulties didn't arise from a lack of understanding of various symbols. The models demonstrated the difference between pivots and linkages, and introduced the graphic symbols that were used in the computer displays to represent pivots and linkages.

One hundred and one undergraduates from Carnegie Mellon served as subjects in the experiment.

Results. The most interesting results arose from an analysis of differences among questions. Specifically, animation improved the ability of lower knowledge subjects to answer questions about the motion of a component or component at a joint that were explicitly mentioned by the text. The improvement therefore, was very local. With the supplemented

description and animation, subjects averaged 4.0 correct answers, significantly better than with the 2.3 correct answers with the regular description, $t(18) = 2.85$, $p < .01$. However, in this condition, the display also controlled the order in which components could be inspected. The effect of animation is also consistent with the claim that lower knowledge subjects are text driven; if the text directs them to evaluate the motion of a specified component, they can use the display animation to "read off" the motion. This also explains why there may be no general effect of animation on lower knowledge subjects. Animation does not provide the more general abstract schema that they may need to construct a better mental model of the device. In contrast, the high ability subjects are able to make some inferences about the motions of components, whether or not the text directs them to do so. Higher knowledge subjects generally performed better on questions that depended making inferences from the diagram, irrespective of whether the text mentioned those specific components. Given that the high ability subjects can make some inferences from the diagram without being directed by the text, it follows that the animation will not be so useful to these subjects.

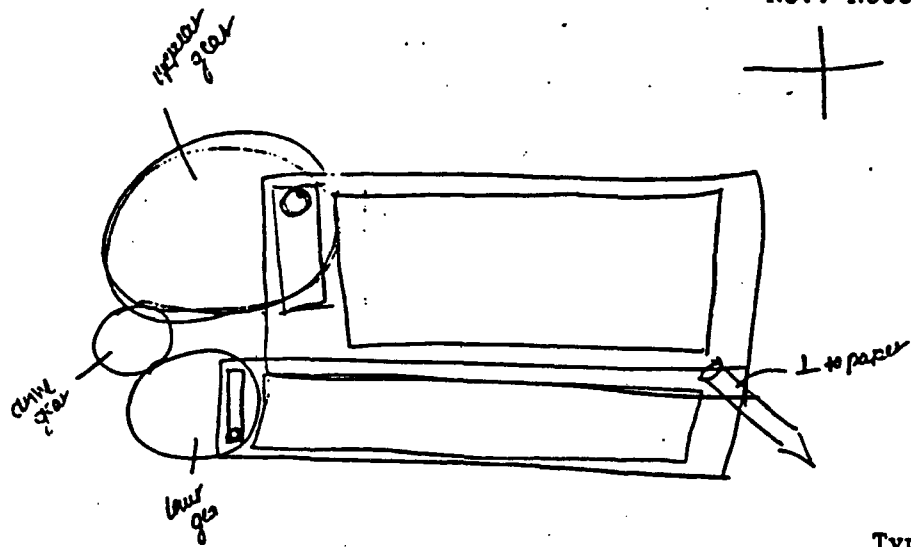
The drawings were scored according to the presence of the major functional components on a scale from 0-7, where 7 points were given to a drawing in which all of the functionally significant structures were present and correctly positioned. No points were given or taken away for quantitative features (such as the number of gear teeth, the size of components) that did not impact on the general functioning of the device. The kinds of drawings were similar to those in Figure 9 and Figure 10 shows the examples of the pencil device.

Insert Figure 10 - Drawings of Pencil Device

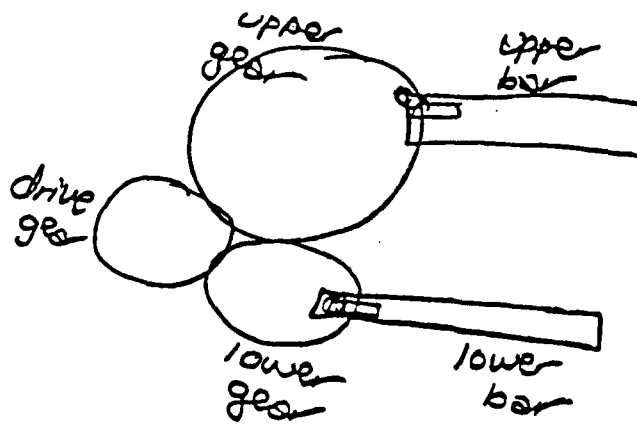
Lower Knowledge Subjects
Comprehension Errors and (Rating of Drawing)

	Display Type		
	Static	Animated	Entire Animation
Supplemented	4.3 (2.0)	3.0 (1.5)	
Normal	3.7 (1.7)	4.7 (0.7)	3.9 (1.7)
Average	4.0 (1.8)	3.8 (1.1)	3.9 (1.7)

In the supplemented-animated condition, the subject did not see the entire display animated, but only joints. Consequently, some of their errors might be attributed to the necessity of integrating pieces of information. However, this hypothesis is not supported, because low mechanical subjects who could animate the entire display had marginally higher error rates, 3.9 errors compared to 3.0 errors in the supplemented-animated condition. In the entirely animated condition, all ten subjects animated the display in the pull cycle and eight also animated it in the push cycle. Thus, all of the information about the motion of various components was available to most of the subjects. Its availability makes it surprising



Typical drawings by
Lower Knowledge
Subjects



Pencil

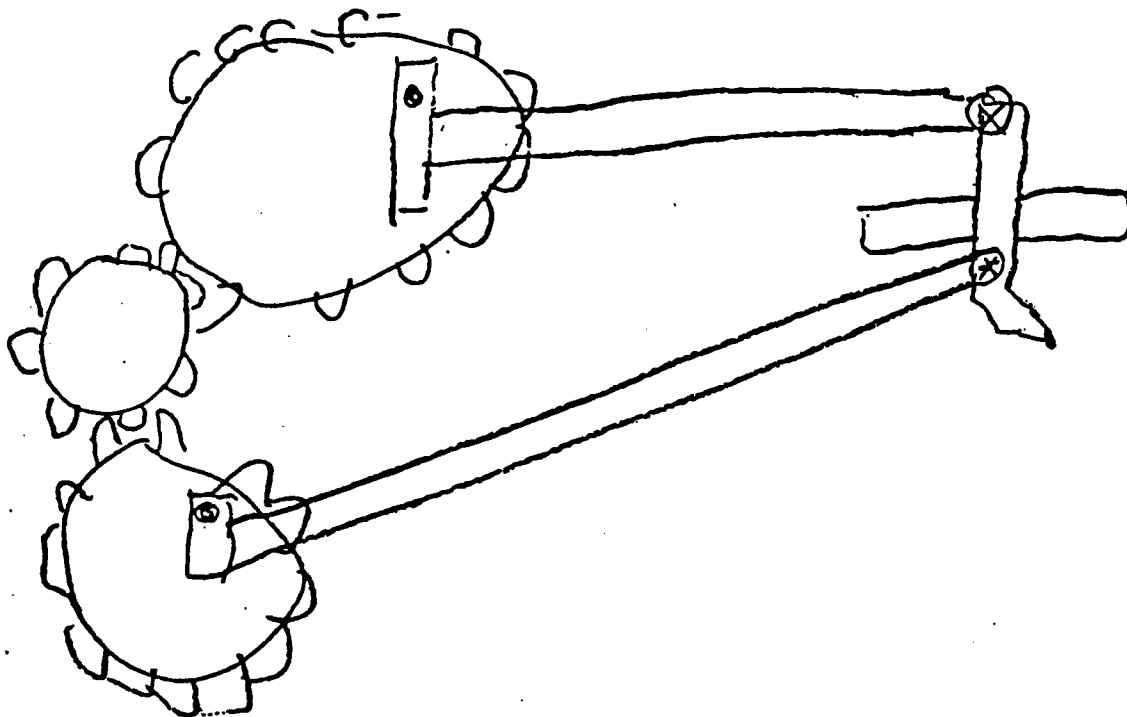


Figure 10

that 40% of these subjects made errors answering the targeted question, namely, what direction does the gearwheel move when the handle is pushed and pulled. The fact that they didn't understand the way motion was transmitted to the gearwheel or the gearwheel's motion itself, even when they could view the entire device, highlights the role encoding plays in understanding the animated display.

The question-answering performance for the static display conditions replicated the results of Experiment 1, showing no improvement due to the supplementedness. Moreover, the results were generally consistent with the hypothesis that low ability subjects have difficulty with mental animation. Subjects in the supplemented condition made a few more errors (4.3 errors) than those in the normal condition (3.7 errors), but the difference was not significant, $t(18) < 1$. Hence, supplementedness in the text alone, without the capability to animate the display, does not help low ability subjects.

The low ability subjects also made consistent, major errors in their drawings the device's structure, suggesting that the low mechanical subjects did not encode or appreciate the relevant geometric structure. The Table above shows the average rating for the low ability subjects' drawings on the scale that ranged from 0 to 7 points. Most drawings, in fact 31 of the 50, had major structural errors in the location, number, and nature of the components (exclusive of the gear teeth) and 38 had major errors in drawing the gear teeth (either no teeth, symmetrical teeth, or teeth that were backwards).

The drawings and question answering were not highly correlated for the low mechanical subjects, $r(48) = .24$, in contrast to the high correlation we will report for the high mechanical ability subjects. The dissociation between the drawing and question answering for the low mechanical subjects suggests that the animated display helped them encode information about the component's movement, but did not improve their understanding of how the motion was determined by the geometric structure of the device.

In contrast to the low mechanical subjects, many high mechanical subjects did understand the structure and motion of the ratchet device, as reflected in significantly better question answering and in their drawings. In fact, better comprehension scores correlated with higher ratings of the drawing across the 51 high mechanical subjects, $r(49) = -0.65$, $p < .01$. An obvious interpretation of this correlation is that an accurate encoding of the structure permitted subjects to make the correct kinematic inferences.

Higher Knowledge Subjects
Comprehension Errors and (Rating of Drawing)

	Type of Diagram		Entire Display
	Static	Animated	
Supplemented	2.3 (5.1)	2.1 (4.5)	
Normal	2.5 (5.3)	1.8 (3.9)	2.4 (4.0)
Average	2.4 (5.2)	2.0 (4.2)	2.4 (4.0)

The average drawing of the high ability subjects included many, but not all, of the major structural components in their proper configuration. Out of a maximum of 7 points, the average rating was 4.5 points, a rating that would typically reflect a drawings that was missing the pivots for the lever and handle, but had a correct representation of the major components, their configuration, and the asymmetry of the gear teeth.

This correlation between the drawing and comprehension score was slightly larger in the static conditions, where subjects had to mentally animate the device, compared to the animated display conditions. One might expect that a correct encoding of the structure would be more crucial to inferring the correct motion in the static conditions. The overall correlation between the question answering and the drawings highlights the important role of selective encoding, both when the display is static and when it is animated. In a cognitive analysis of the components of mechanical ability, we found that one component is knowing what components of a device are mechanically relevant (Hegarty, Just & Morrison, 1988). In this particular task, such knowledge helps one know what is to be coded. For example, it is crucial to the functioning of the ratchet that the teeth be asymmetrical. However, some subjects did not depict them as asymmetrical and the likely interpretation is that they did not code the asymmetry as particularly important. Also, it is crucial to the ratchet device that the lever pivot around a point; but some subjects did not indicate such pivots in their drawing. In general, high mechanical subjects who didn't indicate the functionally important aspects in their drawings also weren't able to answer questions about the motion of various components.

Animated Display Condition				
Description	n	Errors	Time on Diagram (sec)	Time on Text (sec)
Supplemented	7	1.1	166	75
Normal	7	1.0	122	50
Supplemented	3	4.3	104	71
Normal	3	3.7	123	35

Using animation. With both supplemented and normal descriptions, most high knowledge subjects made multiple scans of the upper and lower path and, correspondingly, their error rates were low. For 14 of the 20 subjects, they made an average of 2.2 complete traces of the lower path (which has more components, so that it is easier to identify a trace). The high ability subjects in the normal condition animated fewer times than those in the supplemented condition, but subjects in both conditions usually animated a kinematic pair in the context of scanning along a kinematic chain. The supplemented condition provided structure that the high ability subjects used, but in some sense, may not have needed because they had the strategy of generally following kinematic chains.

Summary. Animation graphics provides a potentially powerful tool for aiding the comprehension of diagrammatic material. What the current research suggests, however, is that animation is not the entire solution. In particular, lower knowledge individuals still need guidance from the text. Moreover, even relatively simple devices appear quite complex to these less knowledgeable individuals who have no schemas to identify the relevant

dimensions and separate them from the irrelevant, but visually complex features. Animation graphics does not necessarily improve their overall comprehension, in spite of clearly eliminating some of the sources of error. In our ongoing research, we are now trying to find out how less knowledgeable individuals or subjects with less spatial ability perceive animated displays.

REFERENCES

- Bennett, G. K. (1969). *Bennett Mechanical Comprehension Test*, New York: The Psychological Corporation.
- Belmont, L., & Marolla, F. A. (1973) Birth order, family size, & intelligence. *Science*, 182, 1096-1101.
- Carpenter, P. A. & Just, M. A. (1989). The role of working memory in language comprehension. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon*. Hillsdale, NJ: Erlbaum.
- Carpenter, P. A., & Just, M. A. (1986). Spatial ability: An information processing approach to psychometrics. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 3). Hillsdale, NJ: Lawrence Erlbaum.
- Carpenter, P. A., & Just, M. A. (1986). Cognitive processes in reading. In J. Orasanu (Ed.), *Reading comprehension: From research to practice*. Hillsdale, NJ: Lawrence Erlbaum.
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404-431.
- Court, J. H., & Raven, J. (1982). *Manual for Raven's Progressive Matrices and Vocabulary Scales* (Research Supplement No. 2, and Part 3, Section 7). London: Lewis & Co.
- Hegarty, M., & Just, M. A. (1989). Understanding machines from text and diagrams. In H. Mandl & J. R. Levin (Eds.), *Knowledge acquisition from text and picture*. North Holland: Elsevier Science Publisher, B.V.
- Hegarty, M., Just, M. A., & Morrison, I. R. (1988). Mental models of mechanical systems: Individual differences in qualitative and quantitative reasoning. *Cognitive Psychology*, 20, 191-236.
- Hegarty, M., Carpenter, P. A., & Just, M. A. (1991). Diagrams in the comprehension of scientific texts. In R. Barr, M. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II). New York: Longman.
- Jensen, A. R. (1987). The *g* beyond factor analysis. In R. R. Ronning, J. A. Glover, J. C. Conoley, & J. C. Witt (Eds.), *The influence of cognitive psychology on testing* (pp. 87-142). Hillsdale, NJ: Erlbaum.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- Just, M. A., & Carpenter, P. A. (1988). Reading and spatial cognition: Reflections from eye fixations. In G. Luer, U. Lass, & J. Shallo-Hoffmann (Eds.), *Eye movement research: Physiological and psychological aspects* (Vol. 2). Lewiston, NY: C. J. Hogrefe.
- Just, M. A., & Carpenter, P. A. (1987). Eye fixations and visual cognition. In G. Luer & U. Lass (Eds.), *Fourth European Conference on Eye Movements* (Vol. 1). Lewiston, NY: C. J. Hogrefe.

- Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, 92, 137-172.
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8, 441-480.
- MacDonald, M. C., & Just, M. A. (1989). Changes in activation levels with negation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 633-642.
- MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, 24, 56-98.
- Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence*, 7, 107-127.
- Raven, J. C. (1962). *Advanced Progressive Matrices, Set II*. London: H. K. Lewis & Co. Distributed in the USA by The Psychological Corporation, San Antonio, Texas.
- Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence, Vol. 2* (pp. 47-103). Hillsdale, NJ: Erlbaum.

PUBLICATIONS ASSOCIATED WITH THE CONTRACT

CHAPTERS:

- Hegarty, M., Carpenter, P. A., & Just, M. A. (1991). Diagrams in the comprehension of scientific texts. In R. Barr, M. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II). New York: Longman.
- Carpenter, P. A., & Just, M. A. (1989). The role of working memory in language comprehension. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon*. Hillsdale, NJ: Erlbaum.
- Hegarty, M., & Just, M. A. (1989). Understanding machines from text and diagrams. In H. Mandl & J. R. Levin (Eds.), *Knowledge acquisition from text and picture*. North Holland: Elsevier Science Publisher, B.V.
- Just, M. A., & Carpenter, P. A. (1988). Reading and spatial cognition: Reflections from eye fixations. In G. Luer, U. Lass, & J. Shallo-Hoffmann (Eds.), *Eye movement research: Physiological and psychological aspects* (Vol. 2). Lewiston, NY: C. J. Hogrefe.

ARTICLES:

- King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30, 580-602.
- Carpenter P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404-431.
- Farah, M. J., Brunn, J. L., Wong, A. B., Wallace, M. A., & Carpenter, P. A. (1990). Frames of reference for allocating attention to space: Evidence from the neglect syndrome. *Neuropsychologia*, 28, 335-347.
- MacDonald, M. C., & Just, M. A. (1989). Changes in activation levels with negation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 633-642.
- Hegarty, M., Just, M. A., & Morrison, I. R. (1988). Mental models of mechanical systems: Individual differences in qualitative and quantitative reasoning. *Cognitive Psychology*, 20, 191-236.

Dr. Phillip L. Ackerman, University of Minnesota, Department of Psychology, 75 East River Road, N218 Elliott Hall, Minneapolis, MN 55455
Dr. David L. Alderton, Code 131, Navy Personnel R&D Center, San Diego, CA 92152-6800
Dr. Terry Allard, Code 1142CS, Office of Naval Research, 800 N. Quincy St., Arlington, VA 22217-5000
Dr. Gregory Anrig, Educational Testing Service, Princeton, NJ 08541
Technical Director, AFRL, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600
Dr. Alan Baddeley, Medical Research Council, Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, ENGLAND
Dr. Patricia Baggett, School of Education, 610 E. University, Rm. 1302D, University of Michigan, Ann Arbor, MI 48109-1259
Dr. Meryl S. Baker, Navy Personnel R&D Center, San Diego, CA 92152-6800
Dr. Pierre Baldi, Div. of Biology, 216-76, Calif. Institute of Technology, Pasadena, CA 91125
Dr. William O. Berry, Director of Life and Environmental Sciences, AFOSR/NL, N1, Bldg. 410, Bolling AFB, DC 20332-6448
Dr. Menucha Brenbaum, Educational Testing Service, Princeton, NJ 08541
Dr. Gautam Blaswas, Dept. of Computer Science, Box 1688, Station B, Vanderbilt University, Nashville, TN 37235
Dr. Robert Breaux, Code 252, Naval Training Systems Ctr., Orlando, FL 32826-3224
Chief, Survey and Market, Analysis Division, Defense Manpower Data Ctr., 1600 Wilson Blvd., #400, Arlington, VA 22209-2593
Dr. Susan Chipman, Cognitive Science Program, Office of Naval Research, 800 North Quincy St., Arlington, VA 22217-5000
Dr. Raymond E. Christal, UES LAMP Science Advisor, AL/HRML, Brooks AFB, TX 78235
Director, Manpower and Training Program, Center for Naval Analyses, 4401 Ford Avenue, P.O. Box 16268, Alexandria, VA 22302-0268
Chief of Naval Education and Training (N-5), NAS Pensacola, FL 32508
Dr. Paul Cobb, Purdue University, Education Building, W. Lafayette, IN 47907
Office of Naval Research, Code 1142, 800 N. Quincy Street, Arlington, VA 22217-5000
Director, Training Systems Department, Code 15A, Navy Personnel R&D Center, San Diego, CA 92152-6800
Chairman, MPT R&D Committee, Office of Naval Research, Code 222, Arlington, VA 22217-5000
Director, Research and Analysis Division, Navy Recruiting Command, Code 223, 4015 Wilson Blvd., Arlington, VA 22217-5000
Library, Code 231, Navy Personnel R&D Center, San Diego, CA 92152
Naval Training Systems Ctr., Attn: Dr. Robert Hays, Code 262, 12350 Research Parkway, Orlando, FL 32826-3224
Chair, Department of Admin. Sciences, Code AS, Naval Postgraduate School, Monterey, CA 93943-5100
Headquarters, United States Marine Corps, Code MA, Washington, DC 20381-0001
Assistant for Planning and Tech. Development, Office of the DCNO, MPT, OP-01B2, Dept. of the Navy, AA-1822, Washington, DC 20350-2000
R&D Coordinator, Attn: Jan Hart, Office of the DCNO, MPT, Op-11K1, Dept. of the Navy, AA-G817, Washington, DC 20370-2000
Deputy Director, Military Personnel Policy Division, Office of the DCNO, MPT, Op-13B, Dept. of the Navy, AA-1825, Washington, DC 20370-2000
Head, Military Compensation, Policy Branch, Office of the DCNO, MPT, Op-134, Dept. of the Navy, AA-2837, Washington, DC 20370-2000
Dr. Stanley Collyer, Office of Naval Technology, Code 222, 800 N. Quincy Street, Arlington, VA 22217-5000
Commanding Officer, Naval Research Laboratory, Code 4627, Washington, DC 20375-5000
Dr. John M. Cornwell, Department of Psychology, I/O Psychology Program, Tulane University, New Orleans, LA 70118
Dr. Michael Cowen, Code 142, Navy Personnel R&D Center, San Diego, CA 92152-6800
Dr. Peter Culicover, Director, Center for Cognitive Science, 208 Ohio Stadium East, 1961 Tuttle Park Place, Columbus, OH 43210-1102
Dr. Diane Damos, Dept. of Human Factors, Institute of Safety & Systems Management, Univ. of Southern California, Los Angeles, CA 90089-0021
Dr. Charles E. Davis, Educational Testing Service, Mail Stop 22-T, Princeton, NJ 08541
Dr. Michael Drillings, Basic Research Office, Army Research Institute, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600
Defense Technical Information Center, DTIC/DDA-2, Cameron Station, Bldg 5, Alexandria, VA 22314
Dr. Richard Duran, Graduate School of Education, University of California, Santa Barbara, CA 93106
Dr. John Ellis, Navy Personnel R&D Center, Code 15, San Diego, CA 92152-6800
ERIC Facility-Acquisitions, 1301 Piccard Drive, Suite 300, Rockville, MD 20850-4305
Dr. Beatrice J. Farr, Army Research Institute, PERI-IC, 5001 Eisenhower Avenue, Alexandria, VA 22333
Dr. Marshall J. Farr, Farr-Sight Co., 2520 North Vernon Street, Arlington, VA 22207
Dr. P.A. Federico, Code 51, NPRDC, San Diego, CA 92152-6800
Dr. Richard L. Ferguson, American College Testing, P.O. Box 168, Iowa City, IA 52243
Dr. Victor Fields, Department of Psychology, Montgomery College, Rockville, MD 20850
Dr. Michael Flanigan, Code 16, NPRDC, San Diego, CA 92152-6800
Dr. J. D. Fletcher, Institute for Defense Analyses, 1801 N. Beauregard St., Alexandria, VA 22311
Dr. Norman Frederiksen, Educational Testing Service (05-R), Princeton, NJ 08541
Dr. Alfred R. Fregly, AFOSR/NL, Bldg. 410, Bolling AFB, DC 20332-6448
Dr. Phillip Gillis, ARI-Fort Gordon, ATTN: PERI-HCD, Fort Gordon, GA 30905
Dr. Herbert Ginsburg, Box 184, Teachers College, Columbia University, 525 West 121st Street, New York, NY 10027
Dr. Drew Gitomer, Educational Testing Service, Princeton, NJ 08541
Mr. Harold Goldstein, University of DC, Dept. Civil Engineering, Bldg. 42, Room 112, 4200 Connecticut Avenue, N.W., Washington, DC 20008
Dr. Sherrie Gott, AFHRL/MOMJ, Brooks AFB, TX 78235-5601
Dr. T. Govindaraj, Georgia Institute of Tech., School of Industrial and Systems Engineering, Atlanta, GA 30332-0205
Dr. Marilyn K. Gowing, Office of Personnel R&D, 1900 E St. N.W., Room 6462, Office of Personnel Management, Washington, DC 20415
Dr. Arthur C. Graesser, Department of Psychology, Memphis State University, Memphis, TN 38152
Dr. Wayne Gray, Graduate School of Education, Fordham University, 113 West 60th Street, New York, NY 10023
Dr. James G. Greeno, School of Education, Stanford University, Room 311, Stanford, CA 94305
Dr. Henry M. Half, Half Resources, Inc., 4918 33rd Road, North, Arlington, VA 22207
Janice Hart, Department of the Navy, Joint CALS Management Office, 5109 Leesburg Pike, Skyline 6, Room 701, Falls Church, VA 22041
Dr. James Hiebert, Department of Educational Development, University of Delaware, Newark, DE 19716
Dr. William Howell, Chief Scientist, AFHRL/CA, Brooks AFB, TX 78235-5601
Dr. Claude Janvier, CIRADE, X-7120, UQAM, P.O. Box 8888, Succ. A, Montreal, Quebec H3C 3P8, CANADA
Dr. Edgar M. Johnson, Technical Director, U.S. Army Research Institute, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600
Dr. Ruth Kanfer, University of Minnesota, Department of Psychology, Elliott Hall, 75 E. River Road, Minneapolis, MN 55455
Dr. Michael Kaplan, Office of Basic Research, U.S. Army Research Institute, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600
Dr. Wendy Kellogg, IBM T. J. Watson Research Ctr., P.O. Box 704, Yorktown Heights, NY 10598
Dr. Patrick Kyllonen, AFHRL/MOEL, Brooks AFB, TX 78235
Richard Lanterman, Commandant (G-PWP), US Coast Guard, 2100 Second St., SW, Washington, DC 20593-0001
COMMANDER USARI, PERI-BR (ATTN: Lawton), 5001 Eisenhower Avenue, Alexandria, VA 22333-5600
Dr. Yuh-Jeng Lee, Department of Computer Science, Code CS/LE, Naval Postgraduate School, Monterey, CA 93943
Dr. Richard Lesh, Educational Testing Service, Princeton, NJ 08541
Dr. Michael Levine, Educational Psychology, 210 Education Bldg., 1310 South Sixth Street, University of IL, Champaign, IL 61820-6990
Dr. Marla C. Linn, Graduate School of Education, EMST, Tolman Hall, University of California, Berkeley, CA 94720
Logicon Inc. (Attn: Library), Tactical and Training Systems Division, P.O. Box 85158, San Diego, CA 92138-5158
LT(N) C. D. F. Lyon, Command Personnel Applied, Research Coordinate, Maritime Command Headquarters, FMO Halifax, N.S., B3K 2X0, CANADA
Vern M. Malec, NPRDC, Code 142, San Diego, CA 92152-6800
Dr. William L. Maloy, Code 04, NETPMSA, Pensacola, FL 32509-5000
Dr. Mary Marino, Directorate, Instructional Technology, HQ USAFA/DFET, USAF Academy, CO 80840-5017
Dr. Sandra P. Marshall, Dept. of Psychology, San Diego State University, San Diego, CA 92182
Dr. Clessen J. Martin, Head, DoD Coordinator, Recruiting Plans and Programs, Branch Code: PERS 2FF/234, Navy Annex, Room 2832, Washington, DC 20350
Dr. John Mayer, Technical Communications Program, 2360 Bonisteel Blvd., University of Michigan, Ann Arbor, MI 48109
Dr. Richard E. Mayer, Department of Psychology, University of California, Santa Barbara, CA 93106
Dr. Joseph McLachlan, Navy Personnel Research and Development Center, Code 14, San Diego, CA 92152-6800
Dr. Michael McNeese, DET-1, ALICFHI, BLDG 248, Wright-Patterson AFB, OH 45432
Dr. Alan Meyrowitz, Naval Research Laboratory, Code 5510, 4555 Overlook Ave., SW, Washington, DC 20375-5000
Dr. E. Muraki, Educational Testing Service, Rosedale Road, Princeton, NJ 08541
Distribution Support Div., Bureau of Naval Personnel, PERS-471D, Washington, DC 20370
Navy Supply Systems Command, NAVSUP 5512, ATTN: Sandra Borden, Washington, DC 20376-5000

Director, Instr. Devel. and Educ. Prog. Spt. Dept., Naval Educ. & Tng. Prog., Mgt. Spt. Activity (NETPMSA), Pensacola, FL 32509-5100
 Head, Leadership Branch, Naval Military Personnel Command, ATTN: LT Gary Kent, NMPC-621, Dept. of the Navy, AA-1803, Washington, DC 20370-5620
 Director, Training Systems Department, NPRDC (Code 14), San Diego, CA 92152-6800
 Director, Training Technology Department, NPRDC (Code 15), San Diego, CA 92152-6800
 Library, NPRDC, Code 041, San Diego, CA 92152-6800
 Librarian, Naval Center for Applied Research in Artificial Intelligence, Naval Research Laboratory, Code 5510, Washington, DC 20375-5000
 Office of Naval Research, Code 1142CS, 800 N. Quincy Street, Arlington, VA 22217-5000
 Special Assistant for Research Management, Chief of Naval Personnel (PERS-01JT), Department of the Navy, Washington, DC 20350-2000
 Dr. Judith Orsenu, Mail Stop 239-1, NASA Ames Research Center, Moffett Field, CA 94035
 Dr. Jesse Orlansky, Institute for Defense Analyses, 1801 N. Beauregard St., Alexandria, VA 22311
 Dr. Okchoon Park, Army Research Institute, PERI-2, 5001 Eisenhower Avenue, Alexandria, VA 22333
 Wayne M. Patience, American Council on Education, GED Testing Service, Suite 20, One Dupont Circle, NW, Washington, DC 20036
 Dr. Roy Pea, Institute for the Learning Sciences, Northwestern University, 1890 Maple Avenue, Evanston, IL 60201
 Dr. Ray S. Perez, ARI (PERI-II), 5001 Eisenhower Avenue, Alexandria, VA 22333
 Dr. Nancy N. Perry, Naval Education and Training Program Support Activity, Code-047, Building 2435, Pensacola, FL 32509-5000
 Dept. of Administrative Sciences, Code 54, Naval Postgraduate School, Monterey, CA 93943-5026
 Dr. Peter Pirolli, School of Education, University of California, Berkeley, CA 94720
 Dr. Martha Polson, Department of Psychology, University of Colorado, Boulder, CO 80309-0344
 Dr. Stephen Reder, NWREL, 101 SW Main, Suite 500, Portland, OR 97204
 Dr. J. Wesley Regian, AFHRL/IDI, Brooks AFB, TX 78235
 Dr. Brian Reiser, Cognitive Science Lab, 221 Nassau Street, Princeton University, Princeton, NJ 08542
 Dr. Gilbert Ricard, Mail Stop K01-14, Grumman Aircraft Systems, Bethpage, NY 11714
 Mr. W. A. Rizzo, Head, Human Factors Division, Naval Training Systems Center, Code 26, 12350 Research Parkway, Orlando, FL 32826-3224
 Dr. Uda G. Roberts, Science, Education and Transportation Program, Office of Tech. Assessment, Congress of the United States, Washington, DC 20510
 Dr. William B. Rouse, Search Technology, Inc., 4725 Peachtree Corners Circle, Suite 200, Norcross, GA 30092
 Dr. Eduardo Sales, Human Factors Division, (Code 262), 12350 Research Parkway, Naval Training Systems Center, Orlando, FL 32826-3224
 Dr. Fumiko Samejima, Department of Psychology, University of Tennessee, 310B Austin Peay Bldg., Knoxville, TN 37966-0900
 Dr. Roger Schank, The Institute for the Learning Sciences, Northwestern University, 1890 Maple Avenue, Evanston, IL 60201
 Dr. Mark Schlager, SRI International, 333 Ravenswood Ave., Room BS-131, Menlo Park, CA 94025
 Dr. Judith W. Segal, OERI, 555 New Jersey Ave., NW, Washington, DC 20208-5648
 Dr. Robert J. Seidel, US Army Research Institute, 5001 Eisenhower Ave., Alexandria, VA 22333
 Mr. Richard J. Shavelson, Graduate School of Education, University of California, Santa Barbara, CA 93106
 Dr. Randall Shumaker, Naval Research Laboratory, Code 5500, 4555 Overlook Avenue, S.W., Washington, DC 20375-5000
 Dr. H. Wallace Sinalo, Manpower Research and Advisory Services, Smithsonian Institution, 801 North Pitt Street, Suite 120, Alexandria, VA 22314-1713
 Dr. Richard E. Snow, School of Education, Stanford University, Stanford, CA 94305
 Dr. Larry Squire, Department of Psychiatry, Vet. Med. Research Foundation, University of California, 3350 La Jolla Village Dr., San Diego, CA 92161
 Dr. Frederick Steinheiser, CIA-ORD, Ames Building, Washington, DC 20505
 Dr. Kurt Steuck, ALHRTI, Brooks AFB, San Antonio, TX 78235-5601
 Dr. Kikumi Tatsuoka, Educational Testing Service, Mail Stop 03-T, Princeton, NJ 08541
 Dr. Douglas Towne, Behavioral Technology Labs, Univ. of S. California, 1228 Spring Street, St. Helena, CA 94574
 Dr. Amos Tversky, Stanford University, Dept. of Psychology, Building 420, Stanford, CA 94305
 Dr. Jerry Vogt, Department of Psychology, St. Norbert College, De Pere, WI 54115-2099
 Dr. Thomas A. Warm, FAA Academy, P.O. Box 25082, Oklahoma City, OK 73125
 Dr. Douglas Wetzel, Code 15, Navy Personnel R&D Center, San Diego, CA 92152-6800
 Dr. Barbara White, School of Education, Tolman Hall, EMST, University of California, Berkeley, CA 94720
 Alexandra K. Wigdor, NRC MH-176, 2101 Constitution Ave., Washington, DC 20418
 Dr. David Wiley, School of Education and Social Policy, Northwestern University, Evanston, IL 60208
 Dr. Kent E. Williams, Department of Industrial and Systems Engineering/MSL, 1900 Kraft Drive, Blacksburg, VA 24060
 Dr. Merlin C. Wittrock, Graduate School of Education, Univ. of Calif., Los Angeles, Los Angeles, CA 90024
 Dr. Joseph L. Young, National Science Foundation, Room 320, 1800 G Street, N.W., Washington, DC 20550
 Director, Human Engineering Lab, Attn: SLCH-E, Aberdeen Proving Grounds, MD 21905-5001
 Tech. Document Ctr., ALHGR-TDC, Wright Patterson AFB, OH 45433-6503
 Dr. Constance W. Atwell, National Eye Inst., SAVP, Bldg 31, Rm. 6A47, 9000 Rockville, Pike, Rockville, MD 20892
 Dr. Kenneth R. Boff, AL/CFH, Wright Patterson AFB, OH 45433-6573
 Dr. Rodney Cocking, NIMH, Basic Behav. & Cog. Sci. Research, 5600 Fishers Lane, Rm. 11C-10, Parklawn Bldg., Rockville, MD 20857
 Nav. Comm., Control & Ocean Surveillance Ctr., RDT & E Division, Attn: Mr. Jeffrey Grossman, Code 44, Bldg. 334, San Diego, CA 92152-5000
 Dr. Daniel B. Jones, U.S. Nuclear Regulatory Commission, NRR/12 E4, Washington, DC 20555
 CDR Robert S. Kennedy, Essex Corporation, 1040 Woodcock Road, Suite 227, Orlando, FL 32803
 LCDR David A. Kobus, Code 23, Nav. Health Resrch Ctr., P.O. Box 85122, San Diego, CA 92186-5122
 Dr. Elizabeth Martin, AL/HRA, Stop 44, Williams AFB, AZ 85240
 Dir., Fleet Liaison Off., NPRDC (Code 01F), San Diego, CA 92152-6800
 Math. Division, Off. of Naval Research, Code 1111, 800 N. Quincy St., Arlington, VA 22217-5000
 Dr. Glenn Osga, NOSC, Code 441, San Diego, CA 92152-6800
 Perceptual Science, Code 1142PS, Off. of Naval Research, Arlington, VA 22217-5000
 CDR Frank C. Patho, Nav. Postgrad. School, Code OR/PE, Monterey, CA 93943
 Elizabeth Wald, Off. of Naval Tech., Code 227, 800 N. Quincy St., Arlington, VA 22217-5000
 Director, Contract Research Dept., Office of Naval Research, (Code 11), Arlington, VA 22217-5000
 Program Manager, Operations Research, (Code 1111MA), Office of Naval Research, Arlington, VA 22217-5000
 Director, Life Sciences, (Code 114), Office of Naval Research, Arlington, VA 22217-5000
 Director, Cognitive & Neural Sciences, (Code 1142), Office of Naval Research, Arlington, VA 22217-5000
 Cognitive Science, (Code 1142CS), Office of Naval Research, Arlington, VA 22217-5000
 CDR. J. S. Hanna, Off. of the Deputy, Asst. Sec. of the Navy, (Manpower), 5D800, The Pentagon, Washington, DC 20350-1000
 Head, Manpower, Personnel and Training Branch, Office of the CNO (Op-813), 4A478, The Pentagon, Washington, DC 20350-1000
 Asst. for Manpower and Training, Office of the CNO (Op-911H), 5D772, The Pentagon, Washington, DC 20350-2000
 Dep. Dir., Total Force Training and Ed. Div., Office of the DCNO (MPT), (Op-11B), Dept. of the Navy, Washington, DC 20370-2000
 Technical Director, Dr. John Silva, Naval Health Research Ctr., P.O. Box 85122, San Diego, CA 92186-5122
 Technical Director, Navy Personnel R&D Ctr., Code 01, San Diego, CA 92152-6800
 Director, Manpower Systems Dept., Code 11, Navy Personnel R&D Ctr., San Diego, CA 92152-6800
 Director, Personnel Systems Dept., Code 12, Navy Personnel R&D Ctr., San Diego, CA 92152-6800
 Director, Testing Systems Dept., Code 13, Navy Personnel R&D Ctr., San Diego, CA 92152-6800
 Director, Defense Personnel Security, Research and Education Ctr., Suite E, Bldg. 455, 99 Pacific Street, Monterey, CA 93940-2481
 Chief Scientist, Air Force (AFHRL), Brooks AFB, TX 78235
 Library, Naval War College, Newport, RI 02940
 Academic Programs & Research Branch, Naval Tech. Training Command, Code N-62, NAS Memphis (75), Millington, TN 30854-5056
 Deputy Director Manpower, Personnel and Training Div., Naval Sea Systems Command, ATTN: Code 04MP 511, Washington, DC 20362